Available online at www.sciencedirect.com
SCIENCE DIRECT®
Vision Research 44 (2004) 1217–1225
Vision Research
www.elsevier.com/locate/visres

Visual search in clutter

Preeti Verghese *, Suzanne P. McKee

Smith Kettlewell Eye Research Institute, 2318 Fillmore Street, San Francisco, CA 94115, USA

Received 31 July 2003; received in revised form 3 December 2003

Abstract

Detecting a target in clutter is particularly difficult because the observer must monitor many potential locations to find the target, and because the clutter itself might mask the target. To investigate whether contemporary models of search can account for visual search in clutter, we measured the detection of an oblique string of five aligned dots presented at an unknown location as a function of noise density. Observers judged which of two 200 ms intervals contained the signal string. At a given density, noise composed of oriented pairs of dots greatly degraded detection compared to random dot noise, especially if the paired noise shared the same orientation as the signal. Increasing the orientation difference between the paired noise and the signal improved detection, as did increasing signal length. We successfully modeled these results with an array of multi-scaled oriented detectors optimally tuned for the signal string. These results indicate that search for these simple patterns in noise is based on competing responses in oriented filters.

Keywords: Search; Masking; Noise; Spatial uncertainty; Orientation

1. Introduction

Numerous psychophysical studies have examined the problem of searching for distinct targets among distractors. In most of these studies, care has been taken to ensure that the elements in the search task are distinct, non-overlapping and presented at equal eccentricity (Eckstein, 1998; Morgan, Ward, & Castet, 1998; Palmer, 1994; Palmer, Ames, & Lindsey, 1993; Verghese & Stone, 1995). However, objects of interest in the real world are rarely presented under such carefully controlled situations. Typically the scene is cluttered and the target might be partially obscured by overlapping items. Take the example of looking for a needle in a haystack. Not only are we uncertain about where the needle is, but its location might be masked by wisps of straw. This task has both the location uncertainty associated with a search task, along with potential masking of the stimulus by overlapping distractors. This is the problem of search in the presence of clutter. Our approach is to devise a cluttered search task and to measure human search performance in such a task. We model the search data with a variant of the signal detection theory models that have been used so far. Rather than dealing with an abstract representation of signal and noise (Burgess & Ghandeharian, 1984; Eckstein & Whiting, 1996; Graham, Kramer, & Yager, 1987; Palmer et al., 1993; Shaw, 1980, 1982; Swensson & Judy, 1981), we will consider the outputs of biologically plausible detectors, followed by a decision rule as described in other studies. Considering the output of realistic detectors has an added advantage; it takes into account the local masking of the target by distractors that happen to fall in the vicinity of the target.

As an approximation to the needle in a haystack problem, we used a target that was an oriented string of dots, and distractors that were random noise (Uttal, 1975). A graphical illustration of the stimulus is shown in Fig. 1a. A target string made up of 5-equally spaced dots oriented 45° counterclockwise can occur at one of the four corners of an invisible square in the target interval. (In this case it appears in the top right corner.) We measured the ability to find the target in clutter when we presented the target in one, two or four locations, and manipulated the discriminability of the target with respect to the distractors. These manipulations included examining the effects of paired noise compared to random noise, varying the orientation of

* Corresponding author. Fax: +1-4153458455.
E-mail address: preeti@ski.org (P. Verghese).

0042-6989/$ - see front matter © 2004 Elsevier Ltd. All rights reserved.
doi:10.1016/j.visres.2003.12.006

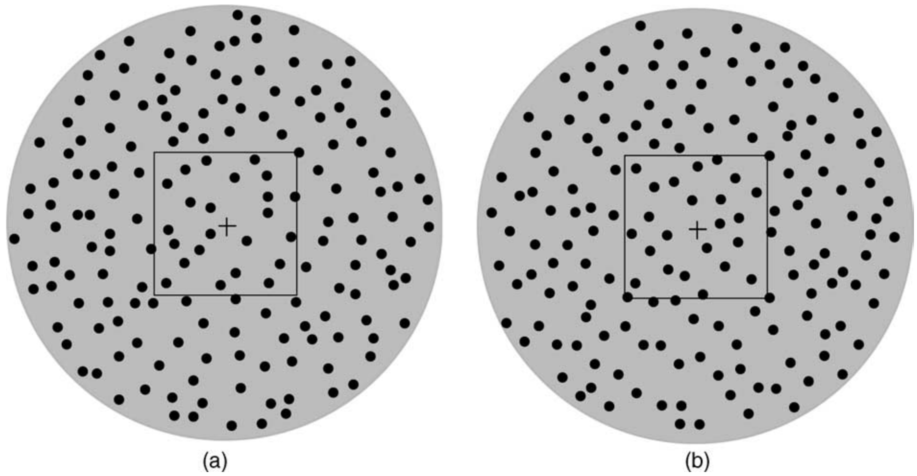

Fig. 1. Example of the target in (a) random noise and (b) in paired noise. The 5-dot target string was oriented 45° counterclockwise with respect to vertical and could occur at one of the fours corners of an invisible square. In both examples, it is at the upper right. The target appeared in one of two intervals and the observer's task was to choose the interval with the target.

the paired noise with respect to the target orientation, and varying the length of the target string. We define "paired noise" as a pair of noise dots with the same spacing as the signal, but with a variable orientation with respect to the signal. In all cases our data are well-fit by a simple model that picks the interval that generates the largest response in an array of orientation-selective detectors.

It is not immediately clear that a decision based on the responses of an array of detectors tuned to the target orientation is sufficient to account for human performance in this task. It seems more likely that the visual system would use a combination of filters selective for the target—filters selective for its spacing as well as its orientation. Of course, the ideal observer would use a template matched exactly to the signal. Classification image studies show that human observers indeed use matched templates for the detection of alphanumeric characters in noise (Watson & Rosenholtz, 1997). While such matched templates are evidently used for over-learned characters such as letters, or when the stimulus matches the properties of simple detectors, other classification image studies reveal a template (filter) that is not exactly matched to the signal. For instance Beard and Ahumada (1999) show that a simple oriented Gabor filter mediates the detection of a vernier stimulus in noise, rather than a template matched exactly to the vernier offset. Our visual search task has a somewhat arbitrarily defined target that is not matched by the known properties of a single class of early filters. We wanted to determine whether search for such a target in the presence of clutter could be explained by an array of simple orientation detectors, or whether a more complex combination of spatial frequency and orientation-tuned detectors was required to predict performance.

2. Methods

The target was an oriented string of dots. There were typically five dots in the string, but this number was varied from 3 to 8 in Experiment 3. The distractors were of two kinds: either random noise, or noise pairs with the same spacing as the target (see Fig. 1). In Experiment 2, the orientation of the noise pairs was varied with respect to the target. For both the signal dots and the paired noise dots, the spacing between the dots was 8′. The target was always oriented 45° counterclockwise with respect to vertical. The noise was randomly distributed throughout the circular display, which was 4° in diameter. In Experiment 1 the target could appear in either one location, one of two locations, or one of four locations. These locations were at the four corners of an imaginary square that was a 2° box centered on fixation. In subsequent experiments, the target could appear anywhere within a 2° square box, centered on fixation. All our experiments employed a 2IFC paradigm with the target + noise in one interval, and the noise alone in the other interval. A new sample of noise was used for each interval. Observers had to choose the interval with the target. Proportion correct was measured as a function of noise density. This was done in separate blocks for each value of noise density and each value of the independent parameter, whether that was number of locations, orientation of the paired noise, or length of the signal string. A trial began with a fixation cross and was followed by the two intervals. Each interval in the trial lasted 200 ms and the blank period between intervals lasted 500 ms.

A total of four observers participated in the experiments. Two of them were naïve as to the purpose of the experiment. The other two observers were the authors.

All four individuals were experienced psychophysical observers.

2.1. Model

The model consisted of oriented detectors that tiled the display area. Since the orientation of the target was known, the model used orientation detectors whose preferred orientation matched that of the target. As for the size of the detector, we compared the performance of the human observer to a model that used detectors of various sizes. A small detector sees little noise but only a fraction of the signal. A large detector sees all of the signal but also a lot of the noise. Detector size was the only free parameter in our model and typically the psychophysical data were best fit by detectors just large enough to cover the signal, as is described in Section 3.

The oriented detector was made up of a pair of Gabor units in sine and cosine phase, respectively. The response of each element of this quadrature pair was squared and summed to calculate a phase-invariant (complex-cell) orientation response. The bandwidth of these units was 1.5 octaves and the aspect ratio was 3, elongated along the axis of preferred orientation. In the case where the target appeared at discrete locations, we assumed that the model monitored the output of detectors at that location. If these detectors were centered on the signal, then the model output was vastly superior to human performance. For the human observer, oculomotor jitter continuously produces small shifts in position. It seems unlikely that there is a perfectly positioned detector at all possible retinal locations, so to make the model more plausible, we randomized the position of the detector to within $\pm\sigma$ (space constant) from the signal.

For the case when the target appeared anywhere in the central 2° we assumed that the model monitored the outputs of the entire array of detectors that tiled this location. Their center-to center spacing of the array of detectors was two times the space constant (standard deviation) of the Gaussian profile of the detector. This spacing allowed a complete coverage of the display area without too much overlap between neighboring filters. This degree of overlap allowed us to assume that the outputs of neighboring detectors were independent. The output of each detector was normalized by the mean response of the detectors in the array and then fed to the decision stage. Because the observer's task was to choose which of two intervals contained the target, the model was designed to choose the largest response in each interval and then pick the interval with the larger of the two chosen responses. As the detectors were matched to the size and orientation of the target, it was quite likely that the detector that saw the target would have the largest response. Errors arose when a detector that saw only noise dots had a stronger response than one that saw the signal. In previous applications of signal detection theory to visual search, the maximum rule was applied to abstract distributions representing the responses to target and distractors. Here, choosing the largest response is in essence the maximum rule applied to the outputs of the oriented detectors, rather than to an idealized response distribution of each element. The maximum rule is close to optimal when there is one signal among a large number of possibilities. Thus, this approach takes into account both the competition from filters responding to noise, and local masking. Because of the high level of added noise, this model assumes that this external noise drowns out any internal noise, so we do not explicitly represent internal noise.

3. Results

3.1. Experiment 1: Target in one of several locations

In this experiment the target appeared in random noise at one location, at one of two locations, or at one of four locations (Fig. 1a). The locations were at the corners of an invisible square and were at an eccentricity of 1.4° from fixation. Trials were blocked by the number of relevant locations. Increasing the number of potential locations is equivalent to increasing the number of distractors, or increasing the set size. The psychometric functions for detecting a target in 1, 2 and 4 locations are shown in Fig. 2a and b for observers NK and PV, respectively. (The upper two graphs are data for the random-noise condition and the lower two graphs are data for the paired-noise condition.) Proportion correct is plotted as a function of the reciprocal density of the noise dots, so as the number of locations is increased, the psychometric functions move towards the right (lower noise densities). The lines are Weibull fits through the data. We summarize these psychometric functions by the threshold, which is the reciprocal noise density that corresponds to 82% correct. The filled symbols in Fig. 3 plot thresholds as a function of the number of possible locations. Thresholds increase with the number of possible locations. The line shows the prediction of the model, which has a detector at each of the relevant locations. The model prediction for finding the target in random noise is a good fit to our observers' data. Note that when the target appears at a single known location, there is no location uncertainty and the only factor that affects the detectability of the target is the presence of the noise dots that fall within the receptive field of the detector at the target location. So the task is similar to detecting an increment on a contrast pedestal or mask. It is important to note that we are using the same model to predict both increment detection when the signal location is known exactly, and 'visual search' when the signal is presented in one of several locations chosen at random.

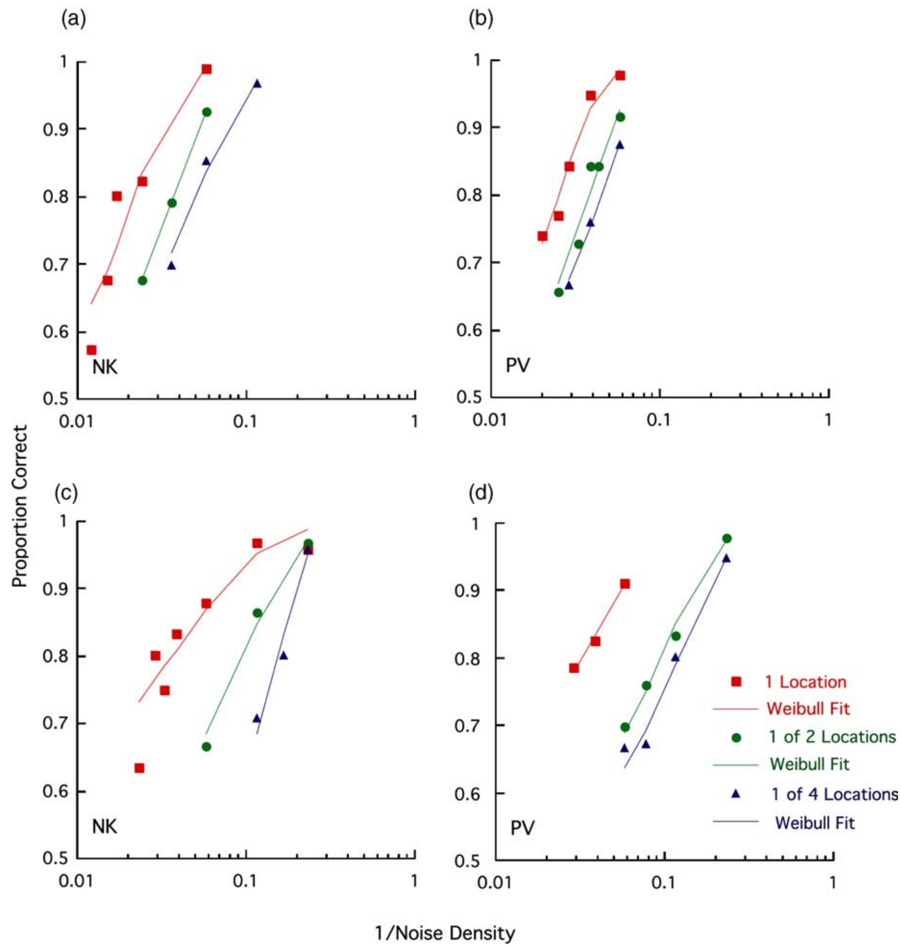

Fig. 2. Visual search for a target in 1 of up to 4 known locations. (a, b) Psychometric functions for detecting the 5-dot target string in random noise for observers NK and PV, respectively. The squares, circles and triangles correspond to the cases when the target was in 1 known location, 1 of 2 locations and 1 of 4 locations. The lines represent Weibull fits through the data. Threshold was taken as the noise density that yielded 82% correct performance. (c, d) Psychometric functions for detecting the target string in paired noise of the same orientation, for observers NK and PV, respectively.

Why does the model mimic human performance? When the observer knows where the signal string will be presented, she monitors a single location and her performance is only limited by the masking effect of the superimposed noise dots. Exactly, the same constraint limits the performance of the model. When the signal is presented at an unknown location, e.g., one of four corners, both the observer and the model examine whatever appears at all four corners of the implicit box in both intervals. From previous studies of visual search, it is known that the performance of real observers is degraded when they are forced to monitor more locations in noise. The reason is simple. Given four locations, there is a greater probability that a random configuration of noise dots at one of the locations will look like the signal string so the real observer will mistakenly choose the noise interval on a higher proportion of trials. Our model will also choose the wrong interval if the noise dots in one of the four locations produces a bigger response in the oriented detectors during the noise interval than the response generated by the signal string (or by a fortuitous noise string) during the signal interval.

Another factor that limits signal detectability in visual search tasks is the similarity between the signal and the noise. Obviously, if the noise is similar to the signal, there is a greater chance that some configuration of the noise will produce a greater response than that generated by the signal. How does a different kind of noise affect detectability for the oriented string of five dots? What happens if the noise consists of pairs of dots with the same orientation and spacing as the signal string? We repeated the previous experiment with the signal in one, two, or four locations with paired noise instead of random noise (Fig. 1b). The psychometric functions for detecting the target string in paired noise are shown in Fig. 2c and d, for observers NK and PV, respectively. For each observer, there is a clear tendency for the psychometric functions to move to the right as the number of locations is increased. The thresholds esti-

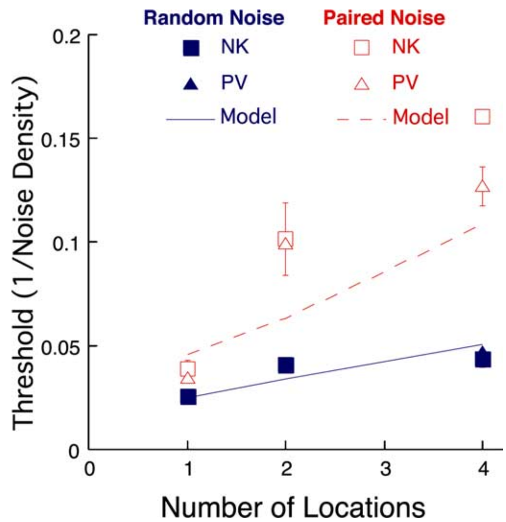

Fig. 3. Threshold noise density as a function of the number of locations. The filled symbols show thresholds in random noise and the open symbols show thresholds in noise pairs that have the same orientation as the signal. The squares and triangles represent data for observers NK and PV, respectively. The error bars represent ±1 standard deviation of the threshold estimate. They are sometimes smaller than the symbols used to plot the data. The solid line is the prediction of our model for search in random noise and the dashed line is the prediction for paired noise. The model fits the data quite well for random noise, but does much better than our observers for paired noise.

mated from these psychometric functions are plotted as open symbols in Fig. 3. Thresholds are higher in paired noise consistent with the subjective impression that the target is harder to see. The dashed line is the model prediction for the paired noise condition. The model predicts the data for a single location with paired noise, but it does significantly better than the two observers for 2 and 4 possible locations. One possible reason for this discrepancy is that two or more sets of paired noise could line up by chance and look very much like the signal, thereby causing the observers to monitor more locations than those specified by the instructions. This can be seen in Fig. 1b, where the paired noise forms several spurious strings that look like the signal. Previous studies have shown that the slope of the psychometric function can be used to estimate the uncertainty, i.e., the number of channels (detectors) that the observer was monitoring (Carrasco, Penpeci-Talgar, & Eckstein, 2000; Green & Swets, 1966; Pelli, 1985; Tyler & Chen, 2000; Verghese & McKee, 2002). As we have indeed measured individual psychometric functions for each of the 1-, 2- and 4-location conditions, we can estimate the number of detectors monitored by the two observers.

We examined the individual psychometric functions (shown in Fig. 2c and d) to determine why observers' thresholds in paired noise rose more steeply than the model prediction. The uncertainty model proposed originally by Green and Swets (1966) (see also Carrasco et al., 2000; Graham et al., 1987; Palmer et al., 1993; Pelli, 1985; Verghese & Stone, 1995) allows one to estimate whether the elevated thresholds for the paired-noise case was because observers monitored more detectors than an ideal observer. We analyzed the psychometric functions using the uncertainty equation of Verghese and McKee (2002) which estimates both the number of detectors used (uncertainty) and the sensitivity of these detectors. These two parameters have different effects on the psychometric function. Increasing the gain shifts the entire psychometric function leftwards to lower values (and lower thresholds), without changing its shape. On the other hand, an increase in uncertainty changes the slope of the psychometric function. The curve is shallow when the observer monitors the single detector that contains the signal, and the curve is steep when the observer monitors many detectors, only one of which contains the signal. Note that if contrast thresholds are specified as a criterion percentage correct, e.g. 82%, they will be lower when the observer is monitoring fewer mechanisms, even if there is no accompanying change in gain.

Fitting individual psychometric functions with the uncertainty models shows that observer NK's thresholds (open squares in Fig. 3) were elevated relative to the model because she monitored more detectors than the number of relevant locations. For NK's data, the uncertainty estimate for set size 2 was 8 times higher than for set size 1, and the estimate for set size 4 was 160 times higher than for set size 1. The changes in uncertainty occurred without accompanying changes in gain—the estimates of detector sensitivity for each of the set size conditions were roughly similar. On the other hand, a similar analysis for observer PV shows that the threshold elevation at set-size 2 and 4 is not due to monitoring more detectors than the number of relevant locations. Instead threshold elevation appears to be due to a diminished sensitivity associated with the set size 2 and 4 conditions, relative to that for set size 1. Perhaps the poor sensitivity with increasing set size is because of a mismatch between the size and/or location of the target and the detectors used by the observers. After all, the target was at the corners of an invisible square. So it is quite possible that the observer was indeed monitoring an incorrect location and/or an inappropriate detector size.

Note that we ran simulations of the model to determine the size of the detector used that best fit the psychophysical performance for each condition. We measured responses to both signal trials and to noise trials for detectors of various sizes. The signal-to-noise ratio (SNR) is an inverted U-shaped function of detector size that reaches a peak at a detector size that roughly spans the signal (see Fig. 4). For random noise, the highest signal-to-noise ratio occurred when the σ of the Gaussian was about 60% of the spacing between the dots. For paired noise, the highest SNR occurred when the space constant of the detector was about 70% of the spacing between dots. However, fitting the data with these detectors of optimal size typically overestimated performance even when the observer was monitoring a single location. The data were better fit by assuming that observers used a detector that was about a factor of two larger.

These results suggest that human observers use a non-optimal filter at eccentric locations. Even though the signals were limited to 1.4° from fixation, human performance for a known location at the fovea is better than for a known location at the periphery. Perhaps, this is because humans tend to use a larger filter at 1.4° eccentricity than at the fovea.

3.2. Experiment 2: Paired distractors of varying orientation

The previous experiment showed that changing the characteristics of the noise affected the detectability of the signal. Clearly, search performance was better for a target embedded in random noise than for a target embedded in paired noise of the same orientation. How does noise of different orientations affect the signal? In this experiment we manipulated the discriminability of the target in noise by varying the orientation of the paired noise with respect to the target. The paired noise had an orientation difference of 0°, 11°, 22°, 45° and 90° with respect to the target. The orientation difference between target and noise was fixed for a block of trials.

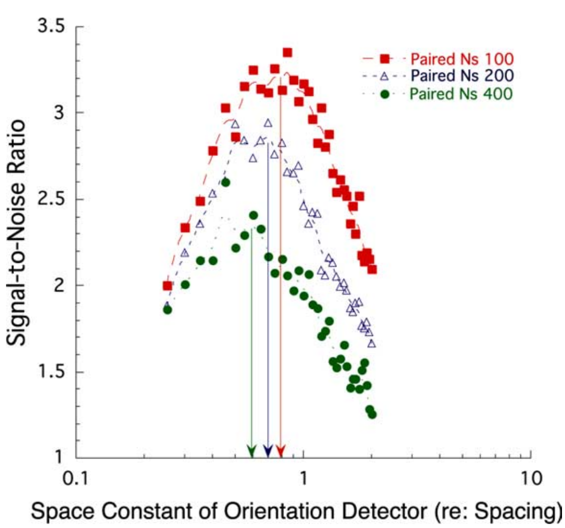

Fig. 4. Signal-to-noise ratio as a function of the size of the orientation detector in our model. At each size, our simulation measured responses to 10,000 trials each of signal present and signal absent trials. The ratio of average signal response to noise response is plotted at three different noise levels—100, 200 and 400 dots. The curves are simply smooth fits through the data points. At each noise level, the arrows represent the detector size that yielded the highest signal-to-noise ratio. There is a tendency for the optimum size to shift to smaller values as the noise level increases.

So far we have dealt with conditions in which the target was in one of several distinct locations. We now relax this constraint so that the target can appear anywhere in the central 2°. Fig. 5 plots the thresholds for our observers as a function of the orientation of the noise pair relative to the target. We modeled this situation by monitoring the output of oriented detectors that tile this region. Thresholds obtained from the individual psychometric functions for each value of target–distractor difference are plotted versus orientation difference for three observers. Thresholds are high when the orientation difference is small but asymptote when the orientation difference reaches 45°. The dashed line plots the model prediction. The model gets the approximate trend of the data. It fits the data well for intermediate values of orientation, although it tends to be slightly worse than human performance at small orientation differences and slightly better at large orientation differences. The model fits can be improved if we use different sizes of detectors for the different orientations of paired noise. Specifically a smaller detector would have predicted the data for small orientation differences better, and a larger detector would have predicted the data for larger orientations. While it is possible that observers used different sizes of detector for different orientations, we chose a fixed pair of detector sizes (see below) to reflect the fixed size of the target.

The entire prediction is based on the assumption that observers are using detectors at two scales. An inspection of the simulated value of signal-to-noise as a function of detector size shows that the optimal size of the detector changes with noise density. Detector size is specified as the space constant of the Gaussian profile of the detector relative to the spacing between the dots. The three curves in Fig. 5 are the ratio of SNR at densities of 100, 200 and 400 dots in the display. It is clear

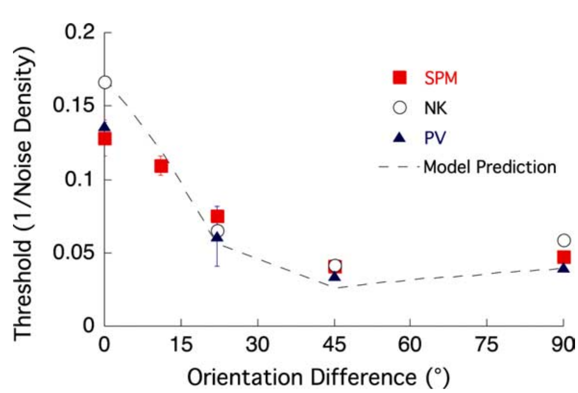

Fig. 5. Threshold noise density as a function of the orientation difference between the target and the paired noise. The different symbols represent data for three observers and the dashed line is the prediction of our model. The model used the same size detector for all values of orientation difference between target and noise.

that the detector size yielding the best signal-to-noise ratio shifts to smaller sizes as the noise level increases. Because noise density was varied within a block of trials to determine thresholds, and because observers could not predict the noise level on an upcoming trial, we assumed that observers used a fixed pair of detector sizes that straddle this preferred size range. In our simulations of the data we used a pair of scales that differed by an octave, i.e. with a space constant of 1.2 and 2.4 times the spacing between the dots. The same set of scales was used for all orientation differences.

3.3. Experiment 3: Signal string length

In this experiment we changed the discriminability of the signal by varying its length. The noise was made up of paired dots of the same spacing and orientation. Fig. 6 plots thresholds for three observers as a function of the length of the string, which varied from 3 dots to 8 dots. The signal is barely detectable when it is a 3-dot string, because it is only one dot longer than the noise pairs. It is easily detected when it is an 8-dot string, because the chance that a random configuration of noise pairs would produce a straight string of eight dots is very small. In this experiment a signal of variable length appeared anywhere in the central 2° in paired noise of the same orientation. (Signal length was fixed in a block of trials.)

Clearly the size of the detector that has the best signal to noise ratio must increase with signal length. We used detectors with space constants 1.05, 1.1, 1.2, 1.5, and 1.8, to fit the psychophysical data for signal lengths of 3, 4, 5, 6, and 8 respectively. The trend is a monotonic increase in detector size with signal length, although the increase is not linear. Keep in mind that the area of our detector increases with its length, so a detector that is long enough to see the eight dots will be large enough to see a lot of the surrounding noise as well. Thus, the best

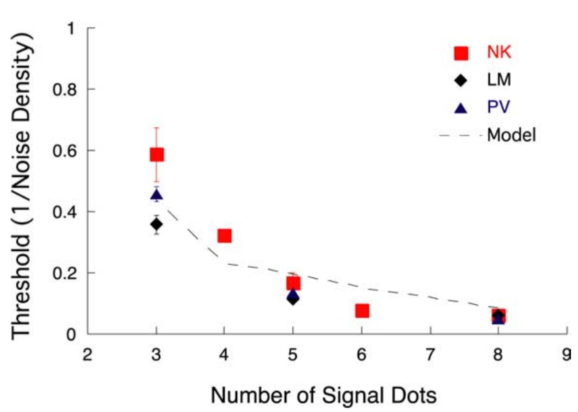

Fig. 6. Threshold noise density as a function of the target length. The different symbols represent data for three observers and the dashed line is the prediction of our model. For each value of target length, the model used the detector size that had the best signal to noise ratio as illustrated in Fig. 5.

SNR will be found for a length that represents a compromise between capturing the whole signal and avoiding too much noise. The model predicts the general trend of the human data, although it underestimates performance slightly for larger signal lengths. This discrepancy favoring the human observer suggests that special neural processing may be used for these longer strings, similar to processes that have been invoked to account for the increased detectability of collinear strings in noise (Field, Hayes, & Hess, 1993; Kovacs & Julesz, 1993). A similar result was obtained for motion trajectories (Verghese, Watamaniuk, McKee, & Grzywacz, 1999) While the detectability of a short signal traveling along a straight path in noise was predicted by the outputs of independent motion detectors, the detectability of longer trajectories far exceeded the prediction of a model based on the output of independent detectors. In the motion domain, it appears that the enhanced detectability of long trajectories is because the first part of the trajectory acts as an implicit cue to subsequent parts (Verghese & McKee, 2002).

4. Discussion

This simple model based on plausible orientation-selective detectors successfully predicts the trend of the data in a cluttered search experiment. It shows that the probability of finding the target decreases as the number of locations monitored increases (the set size effect). Other studies have shown that signal detection theory based on an abstract representation of signal and noise accounts for simple visual search (Eckstein, 1998; Palmer et al., 1993, 2000; Verghese & Stone, 1995). The relevance of this study is that it moves away from the abstract representation to a more realistic representation based on the output of early detectors selective to the signal. Although this is among the first studies to start with the search display itself and to apply biologically plausible filters to explain visual search performance, similar approaches have been adopted in medical imaging (e.g. Abbey & Barrett, 2001; Burgess, Jacobsen, & Judy, 2001; Eckstein, Bochud, & Abbey, 2000). In these studies various filters ranging from a matched template modulated by the human contrast sensitivity function to a combination of spatial frequency and orientation-tuned detectors were used to predict human ability to detect targets in noisy radiographic images.

The studies that have applied signal detection theory to predict visual search accuracy have typically assumed that the abstract distribution of signal and noise distributions is Gaussian. However, simulations based on 10,000 trials show that the response distribution of orientation detectors to both the signal and noise intervals is distinctly non-Gaussian. The distribution has pronounced tails that extends toward larger responses.

At any given noise density, the skew is more evident in signal than in noise intervals, which makes the maximum rule a very effective decision rule.

In our experiments the spacing between the dots was the same for the signal and the paired noise, so that the noise would mimic the signal. The model that we have proposed does not explicitly take the spacing of the dots into account. In fact, the model predictions do not change when this spacing is jittered by a small amount. This is quite different from the predictions of an ideal observer that would certainly take advantage of the fact that the dots in the signal (and noise pairs) were spaced at regular intervals. While the human observer might make use of dot spacing at very low densities, the regular spacing of the signal dots is not so apparent at moderate to high noise densities at which threshold is typically measured. This is because the noise dots often appear in between the signal dots at moderate to high noise densities and disrupt their regular spacing. This might explain why the model does a good job of predicting threshold performance, even when it does not incorporate the regular spacing between signal elements.

There are small discrepancies between the model's predictions and the performance of human observers. A closer comparison of the model fit to the data shows that human performance in paired noise degrades faster with increasing set size than the model prediction (open symbols in Fig. 3). This discrepancy may be due to the fact that the model does not take eccentricity into account, whereas human observers show a large eccentricity effect. Even though the signals were limited to being within 1.4° from fixation, human performance for a known location at the fovea is better than for a known location at an eccentricity of 1.4°. An analysis of the psychometric functions for each set size condition suggests that the poor performance at these eccentric locations is because human observers monitor more locations or use inappropriate detector sizes and locations.

The model also does a reasonable job of predicting search performance as a function of the orientation difference between the target and the distractor pairs (Fig. 5), as well as the difference in length (number of dots in target string) between the target and distractors (Fig. 6). It did underestimate human performance when the orientation difference between target and distractors was small, less than 5°. But given its simplicity (for example, individual detectors were only tuned for the orientation of the target and not for the spacing of the dots in the string), it is impressive how well it predicts human performance. Our results indicate that the signal detection approach, combined with biologically plausible detectors, can predict performance in more realistic search tasks with crowded displays. They also suggest that a more elaborate template that matches the characteristics of the target exactly, such as the size and exact spacing of the dots is not required.

Of course the visual system is more complicated than a front end of oriented detectors followed by a decision stage. But the fact that we can predict general trends in search performance in the presence of local masking demonstrates the potential of this simple model. Note that in the orientation tuning experiments the signal string could differ from the noise in both length and in orientation. We did not have to make up any complicated search rules: search performance is largely captured by a decision process acting on the output of oriented detectors. These results indicate that search for simple patterns in noise is based on competing responses in oriented filters.

We have considered a very simple example of visual search in clutter. Can this approach be extended to visual search in natural scenes where the differences between target and clutter are not easily characterized by the responses to oriented filters? While a more complex target might need a more sophisticated template, there are other factors such as grouping and segmentation that might aid visual search performance under natural conditions. In fact our data hint at such a grouping process. For instance, integration along the signal string might explain why human performance for long signal strings exceeds the prediction of a model based on independent detectors. Nevertheless, a simple extension of search models using the outputs of biologically plausible filters largely accounts for the detectability of a simple target in clutter.

Acknowledgements

This work was supported by NEI grants EY12038 to PV and EY06644 to SPM.

References

Abbey, C. A., & Barrett, H. H. (2001). Human- and model-observer performance in ramp-spectrum noise; effects of regularization and object variability. *Journal of the Optical Society of America A, 18*, 473–488.

Beard, B. L., & Ahumada, A. J. (1999). Detection in fixed and random noise in foveal and parafoveal vision explained by template learning. *Journal of the Optical Society of America A, 16*, 755–763.

Burgess, A. E., & Ghandeharian, H. (1984). Visual signal detection II. Signal–location identification. *Journal of the Optical Society of America A, 1*, 906–912.

Burgess, A. F., Jacobsen, F. L., & Judy, P. F. (2001). Human observer detection experiments with mammograms and power-law noise. *Medical Physics, 28*, 419–437.

Carrasco, M., Penpeci-Talgar, C., & Eckstein, M. P. (2000). Spatial covert attention increases contrast sensitivity across the CSF: support for signal enhancement. *Vision Research, 40*(10–12), 1203–1215.

Eckstein, M. P. (1998). The lower visual search efficiency for conjunctions is due to noise and not serial attentional processing. *Psychological Science, 9*, 111–117.

Eckstein, M. P., & Whiting, J. S. (1996). Visual signal detection in structured backgrounds. I. Effect of number of spatial locations and signal contrast. *Journal of the Optical Society of America A, 13*, 1777–1787.

Eckstein, M. P., Bochud, F. O., & Abbey, C. A. (2000). Visual signal detection in structured backgrounds. IV. Figures of merit for model observers in multiple forced choice with response correlations. *Journal of the Optical Society of America A, 17*, 206–217.

Field, D. J., Hayes, A., & Hess, R. F. (1993). Contour integration by the human visual system: evidence for a local association field. *Vision Research, 33*, 173–193.

Graham, N., Kramer, P., & Yager, D. (1987). Signal-detection models for multidimensional stimuli: probability distributions and combination rules. *Journal of Mathematical Psychology, 31*, 366–409.

Green, D. M., & Swets, J. A. (1966). *Signal detection theory and psychophysics*. New York: John Wiley & Sons.

Kovacs, I., & Julesz, B. (1993). A closed curve is much more than an incomplete one: effect of closure in figure-ground segmentation. *Proceedings of the National Academy of Science of the USA, 90*, 7495–7497.

Morgan, M. J., Ward, R. M., & Castet, E. (1998). Visual search for a tilted target: tests of spatial uncertainty models. *Quarterly Journal of Experimental Psychology A, 51*(2), 347–370.

Palmer, J. (1994). Set-size effects in visual search: the effect of attention is independent of the stimulus for simple tasks. *Vision Research, 34*(13), 1703–1721.

Palmer, J., Ames, C. T., & Lindsey, D. T. (1993). Measuring the effect of attention on simple visual search. *Journal of Experimental Psychology: Human Perception and Performance, 19*(1), 108–130.

Palmer, J., Verghese, P., & Pavel, M. (2000). The psychophysics of visual search. *Vision Research, 40*, 1227–1268.

Pelli, D. G. (1985). Uncertainty explains many aspects of visual contrast detection and discrimination. *Journal of the Optical Society of America A, 2*(9), 1508–1532.

Shaw, M. L. (1980). Identifying attentional and decision-making components in information processing. *Attention and Performance* (vol. VIII). R.S. Nickerson, pp. 277–296.

Shaw, M. L. (1982). Attending to multiple sources of information: I. The integration of information in decision making. *Cognitive Psychology, 14*, 353–409.

Swensson, R. G., & Judy, P. F. (1981). Detection of noisy visual targets: model for the effects of spatial uncertainty and signal to noise ration. *Perception & Psychophysics, 29*, 521–534.

Tyler, C. W., & Chen, C. C. (2000). Signal detection theory in the 2AFC paradigm: attention, channel uncertainty and probability summation. *Vision Research, 40*, 3121–3144.

Uttal, W. R. (1975). *An autocorrelation theory of form detection*. New Jersey: Lawrence Erlbaum.

Verghese, P., & McKee, S. P. (2002). Predicting future motion. *Journal of Vision, 2*(5), 413–423.

Verghese, P., & Stone, L. S. (1995). Combining speed information across space. *Vision Research, 35*, 2811–2823.

Verghese, P., Watamaniuk, S. N., McKee, S. P., & Grzywacz, N. M. (1999). Local motion detectors cannot account for the detectability of an extended trajectory in noise. *Vision Research, 39*(1), 19–30.

Watson, A. B., & Rosenholtz, R. (1997). A Rorschach test for visual classification strategies. Investigative Ophthalmology and Visual Science Abstracts, S1.